Discovering Rare Behaviors of Stochastic Models using Decision Procedures

Arup K. Ghosh, University of Central Florida
Emily Sassano, University of Central Florida
Sumit K. Jha, University of Central Florida
Christopher J. Langmead, Carnegie Mellon Univ
Susmit Jha, UC Berkeley

Outline

- Introduction
- Problem definition
- State of the art
- Our solution
  - Basic idea
  - Algorithm
  - Results
- Conclusion and Future Work

Introduction

- Discovering rare behaviours of stochastic models is an unsolved challenge.

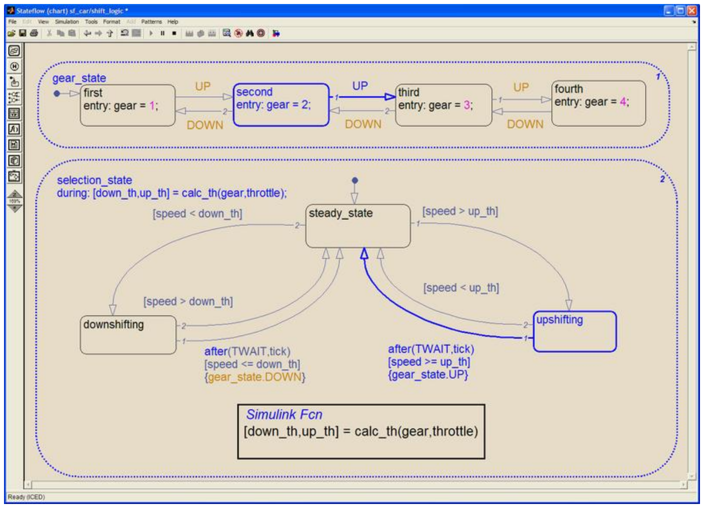
Space Vehicle

Artificial Pancreas

Infectious Disease

- Several Domains:
  - Biomedical Devices,
  - Intelligent Automobiles,
  - Autonomous Robots

Why Stochastic Models? - I



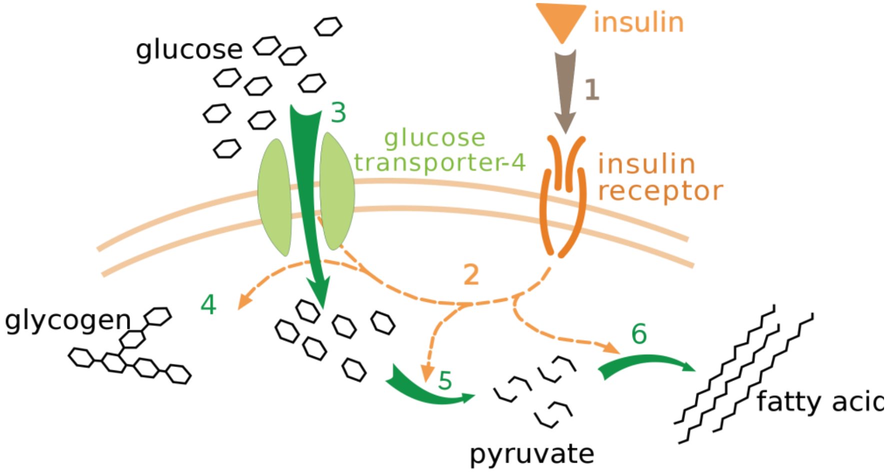
Model of Embedded System

Unpredictable environment

- Even when models are deterministic
- The environment is quite complex e.g. road conditions
  - Stochastic models capture the unpredictability of inputs to the embedded system.

Why Stochastic Models? - II

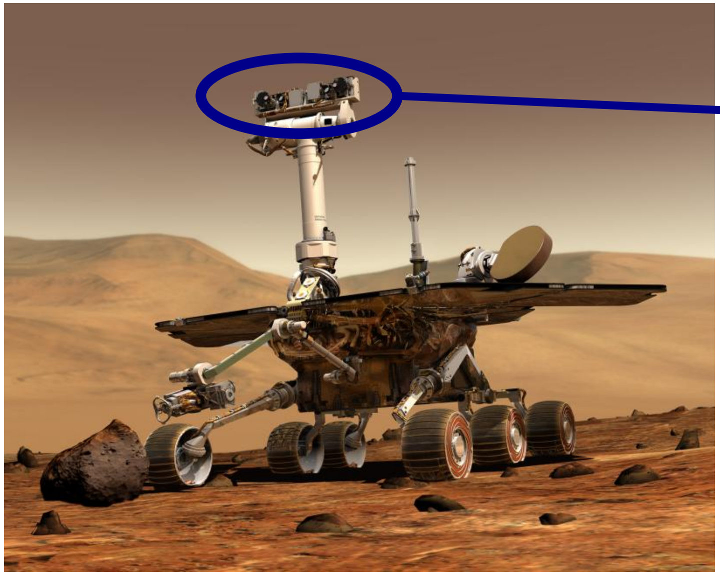
Artificial Pancreas



Insulin Metabolism

- Complex life-critical embedded systems
- Must be proved correct when interacting with fundamentally stochastic systems
  - Biochemical environment models are fundamentally stochastic

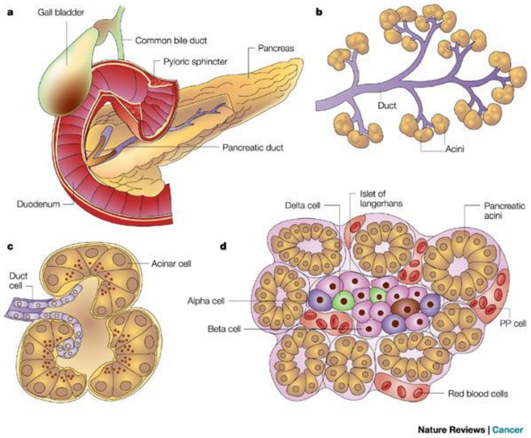
Why Stochastic Models? - III



Stereoscopic Camera for computer vision

- Models of Complex “Intelligent” Software are inherently stochastic
- E.g. Computer Vision algorithms, Machine Learning, Voice Recognition, Feature Extraction

Why Study Stochastic Models ?



- 32,000 Americans diagnosed with Pancreatic Cancer yearly.
- Almost all of them will die!
- Bits and pieces of stochastic biochemical models available!

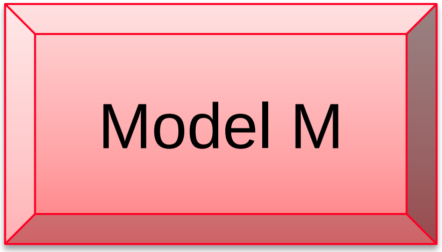
- Next generation vehicles (will) do lane tracking, etc.
- Modify user inputs “if necessary” ?
- Stochastic models of human behavior!

- In 2007, Lehman made $4 billion in profits.
- It lost $3 billion in the spring and then another $4 billion in the following summer.

Why Study Rare Behaviors ?

- Rare but interesting or important behaviors
  - formation of a tumor,
  - spreading of infectious diseases,
  - failure of cyber-physical system

Stochastic Model Simulation



Stochastic Model :

- Naturally equipped with a well defined probability space.
- Example: DTMCs, CTMCs, SDEs
  - Deterministic Simulink Models with probabilistic inputs
  - Probabilistic Simulink Models
- Not an example:
  - Markov Decision Processes
  - C Programs
  - Digital Circuits

SDE Models

- The form of a typical SDE:
  - $dX = b(t, X)\, dt + v(t, X)\, dW$

  where
    - X is a system variable
    - b is Riemann integrable function
    - v is Ito integrable
    - W is Brownian Motion

Why SDE models?

- Model dynamics of complex “less understood” systems.

- Investigation of biological phenomena and cyber-physical systems
  - sensitive to stochastic effects.

Stochastic Model Verification

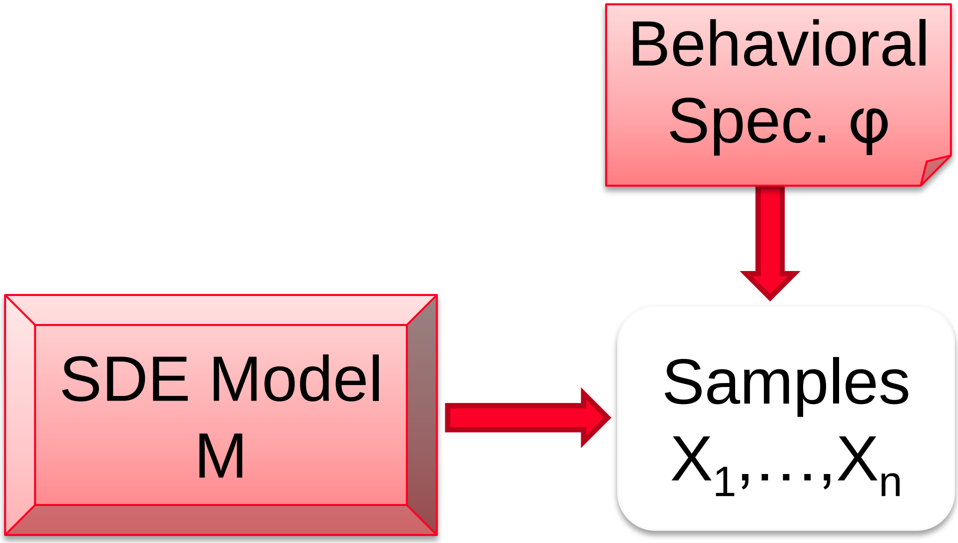
Behavioral Spec. φ

SDE Model M

Behavioral Specification φ should be decidable on a finite sample trace of the model M.

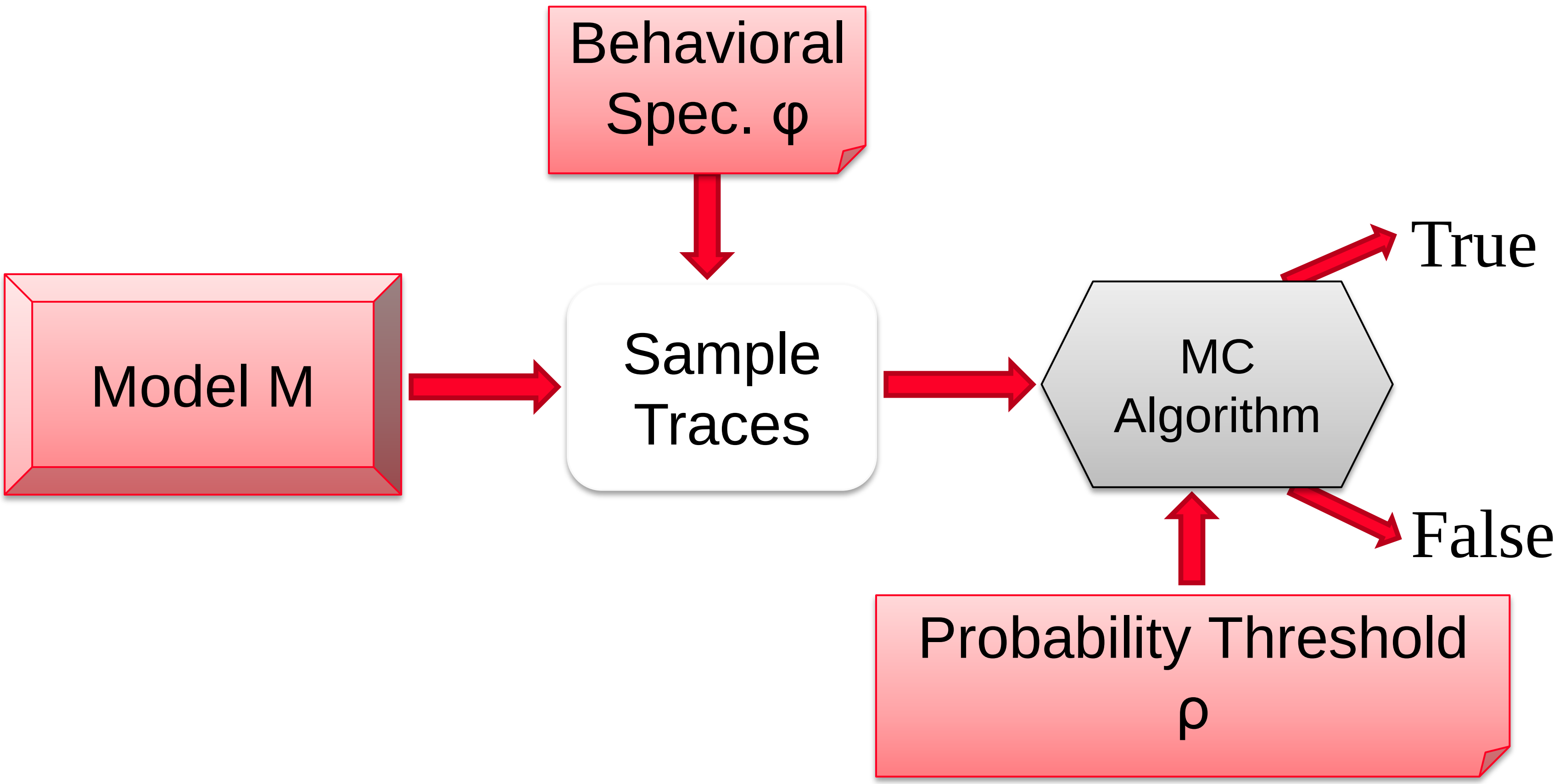
- Example:
  - Bounded Linear Temporal Logic
  - Finite Sate Machine Specifications
  - “Sun will rise in the east within 24 hours”
- Not an example:
  - “Eventually, the sun will rise in the west some day.”

Statistical Model Verification



- Samples X_i should be independent and identically distributed (i.i.d.) according to the model M.
  - $X_i = 1$ if the sample satisfies specification φ; 0 otherwise.
- Example:
  - i.i.d. random sampling of Simulink models
- Not an example:
  - Rare event simulation.
  - Symbolic Testing Strategies

Statistical Model Verification



State of the Art: Statistical

- **Statistical Estimation:**
  - Given samples from a stochastic system,
  - Compute the probability that a system satisfies a given property.
- **Statistical Hypothesis Testing:**
  - Given a required confidence in the answer and a threshold probability
  - Draw as many samples as needed to decide
    - whether a stochastic system satisfies a given property with at least a threshold probability?
- **Both approaches are really useless for rare but interesting behaviors**

State of the Art: Statistical

- Use Girsanov's theorem to change probability measures and then use statistical sampling to explore rare behaviors

- Sumit K. Jha and Christopher Langmead

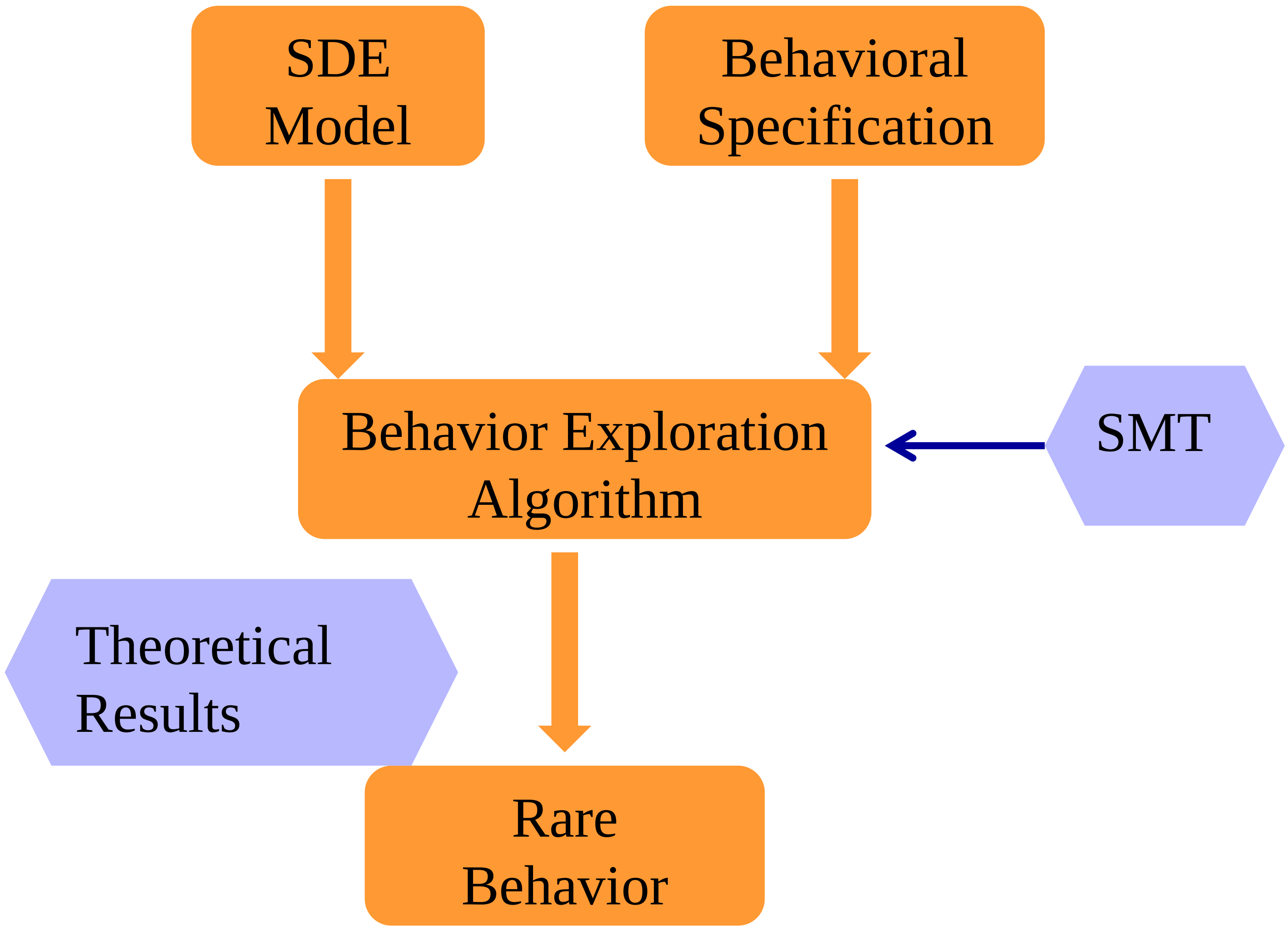
  Understanding rare behaviors in stochastic biological models

  Best Paper at IEEE International Conference on Computational Advances in Bio and Biomedical systems

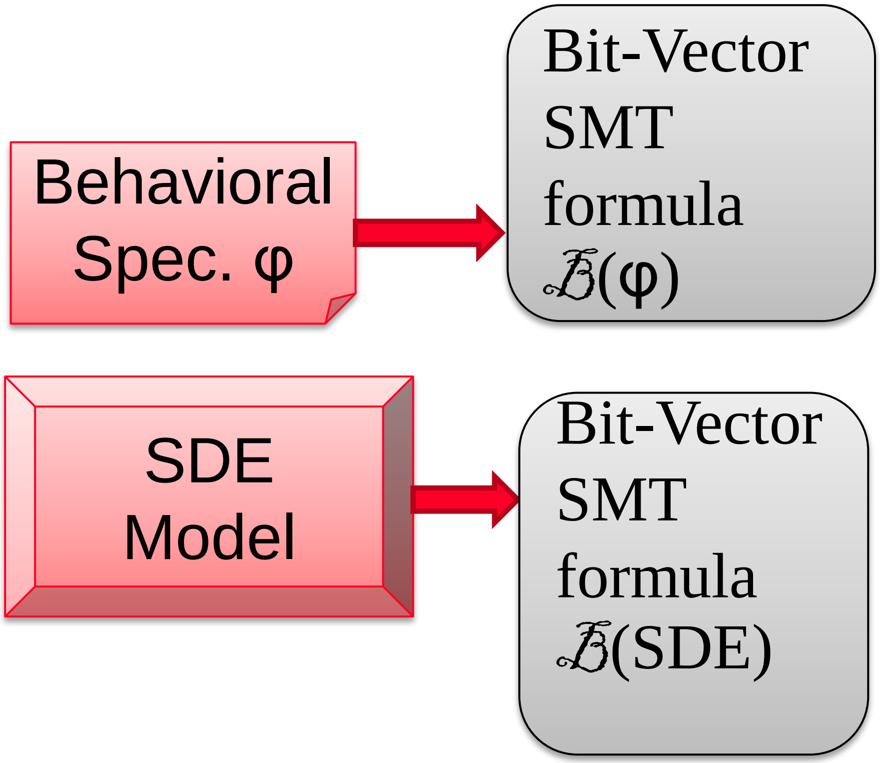
  Journal Version: BMC Bioinformatics

- Tries to explore rare behaviors but not a given rare behavior

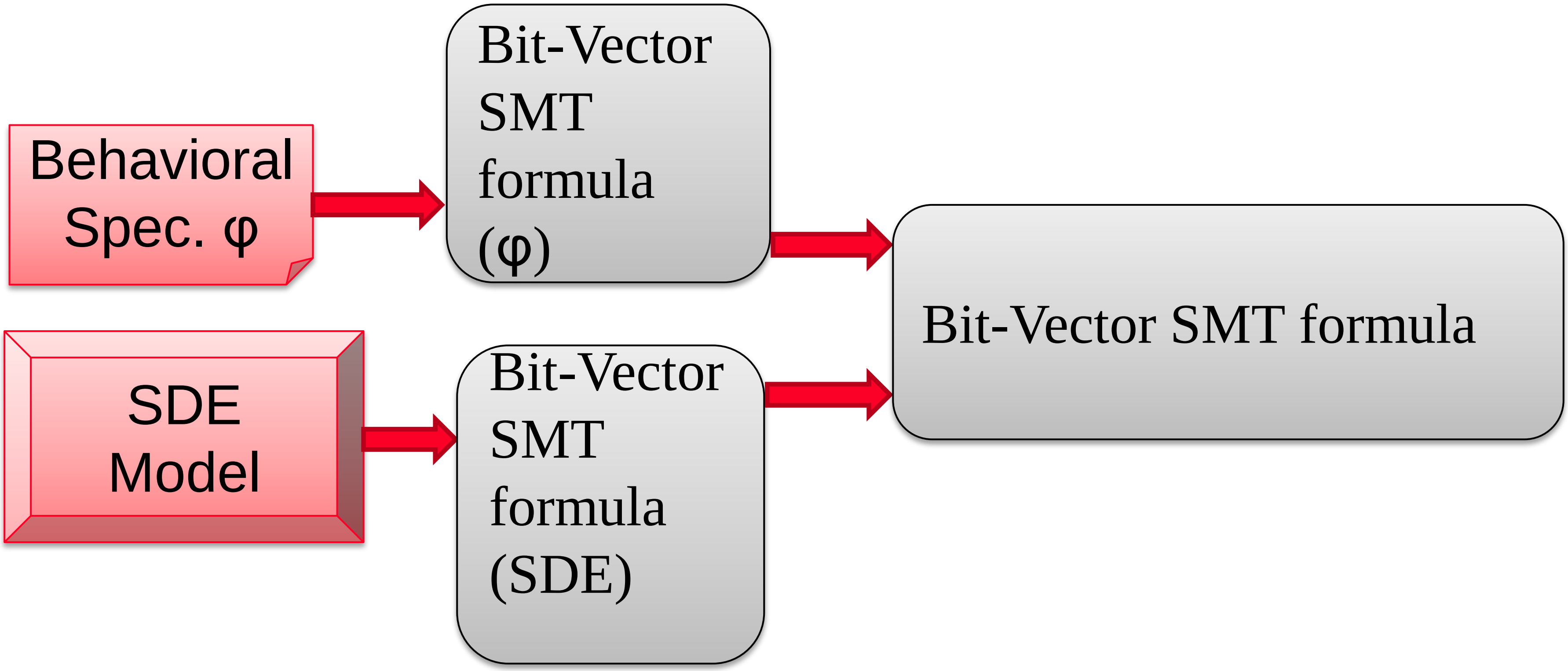
Our New Approach - I



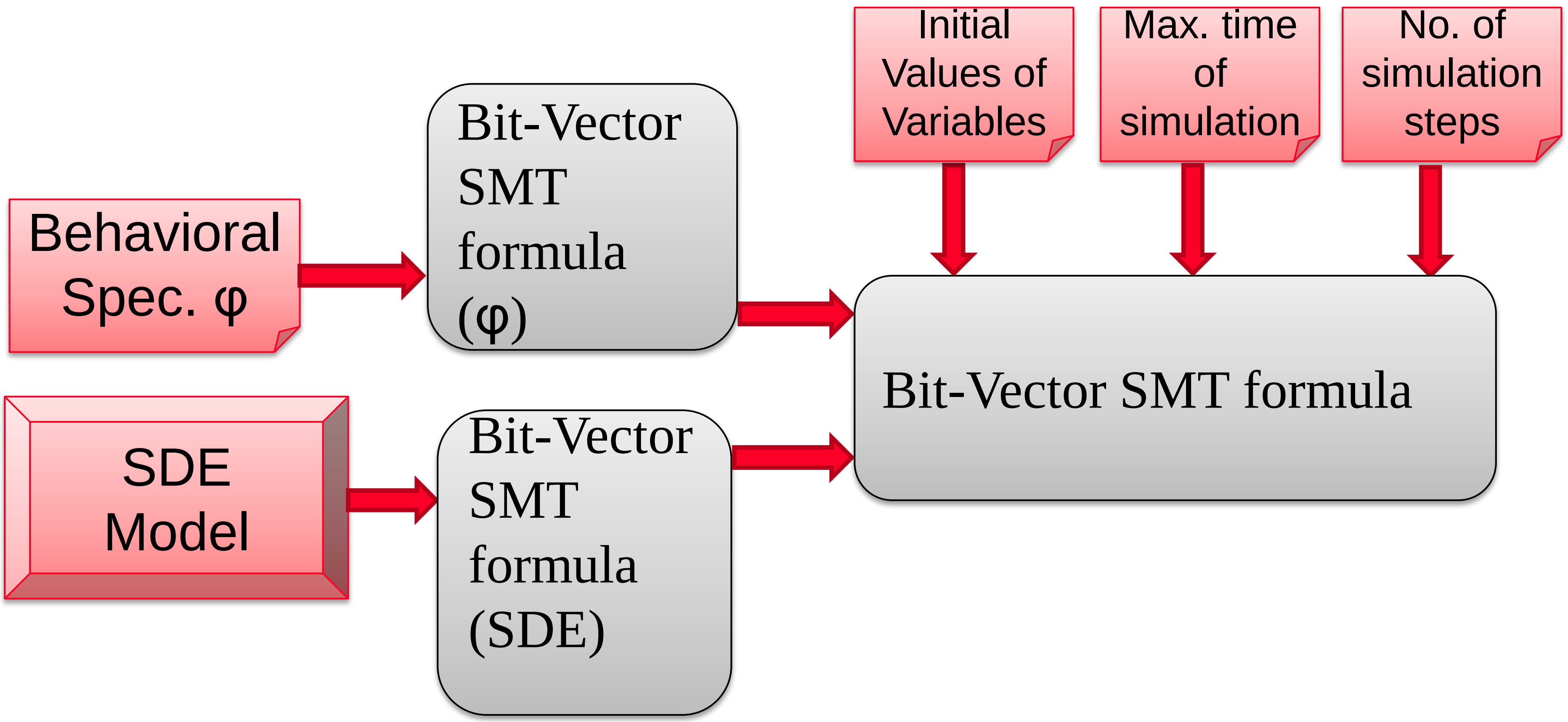
Our Approach – II (RESERCHE)



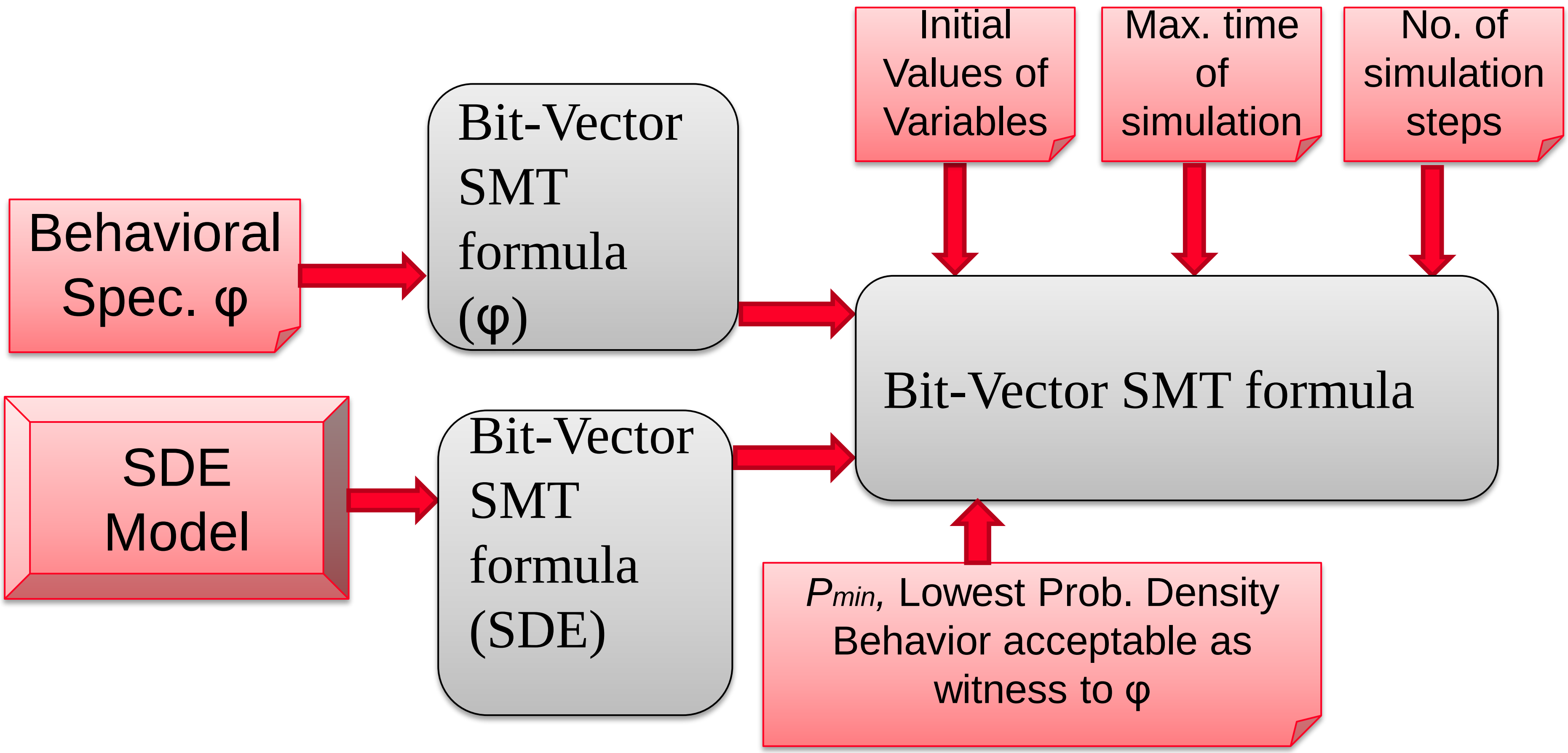
Our Approach – II (RESERCHE)



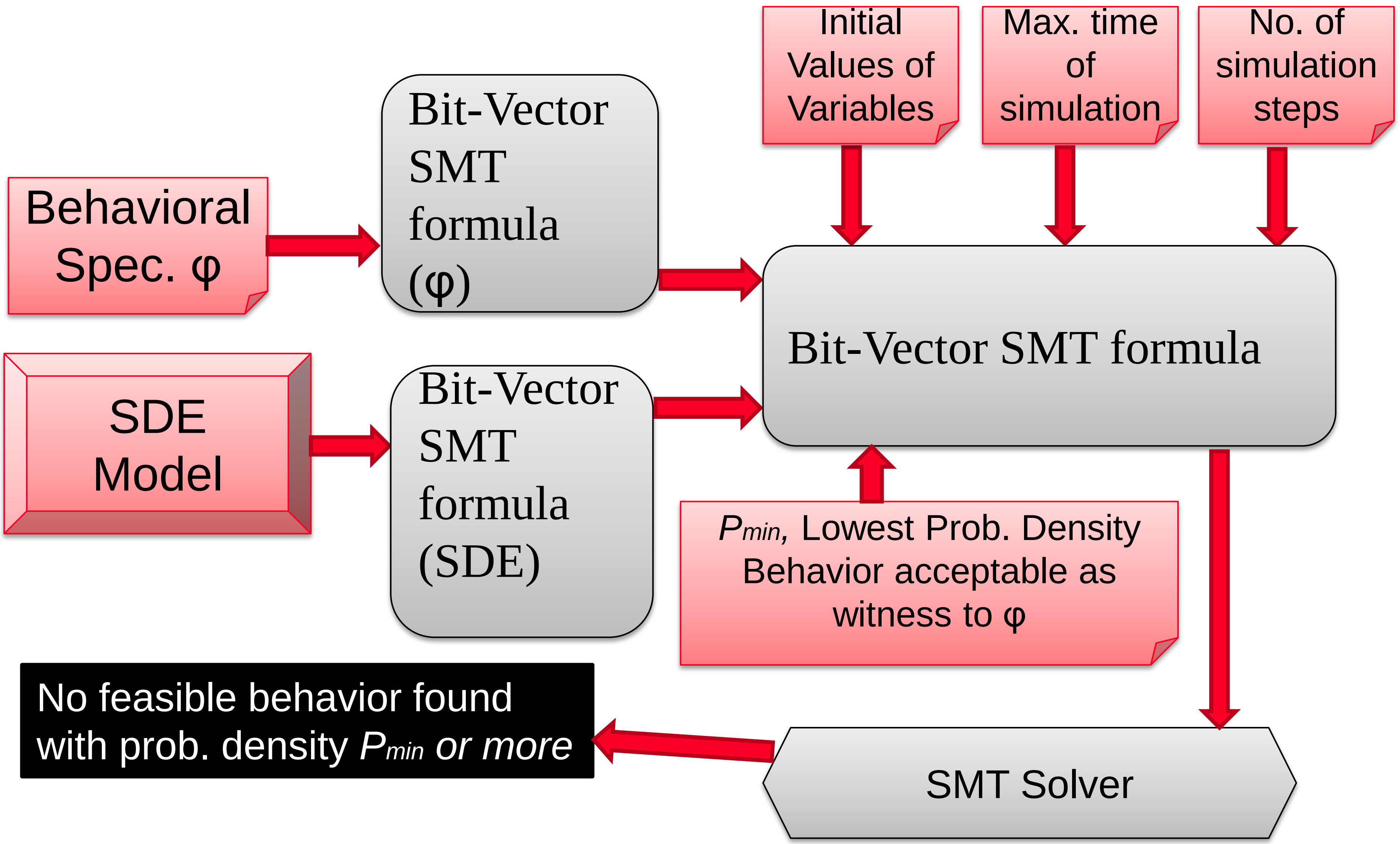
Our Approach – II (RESERCHE)



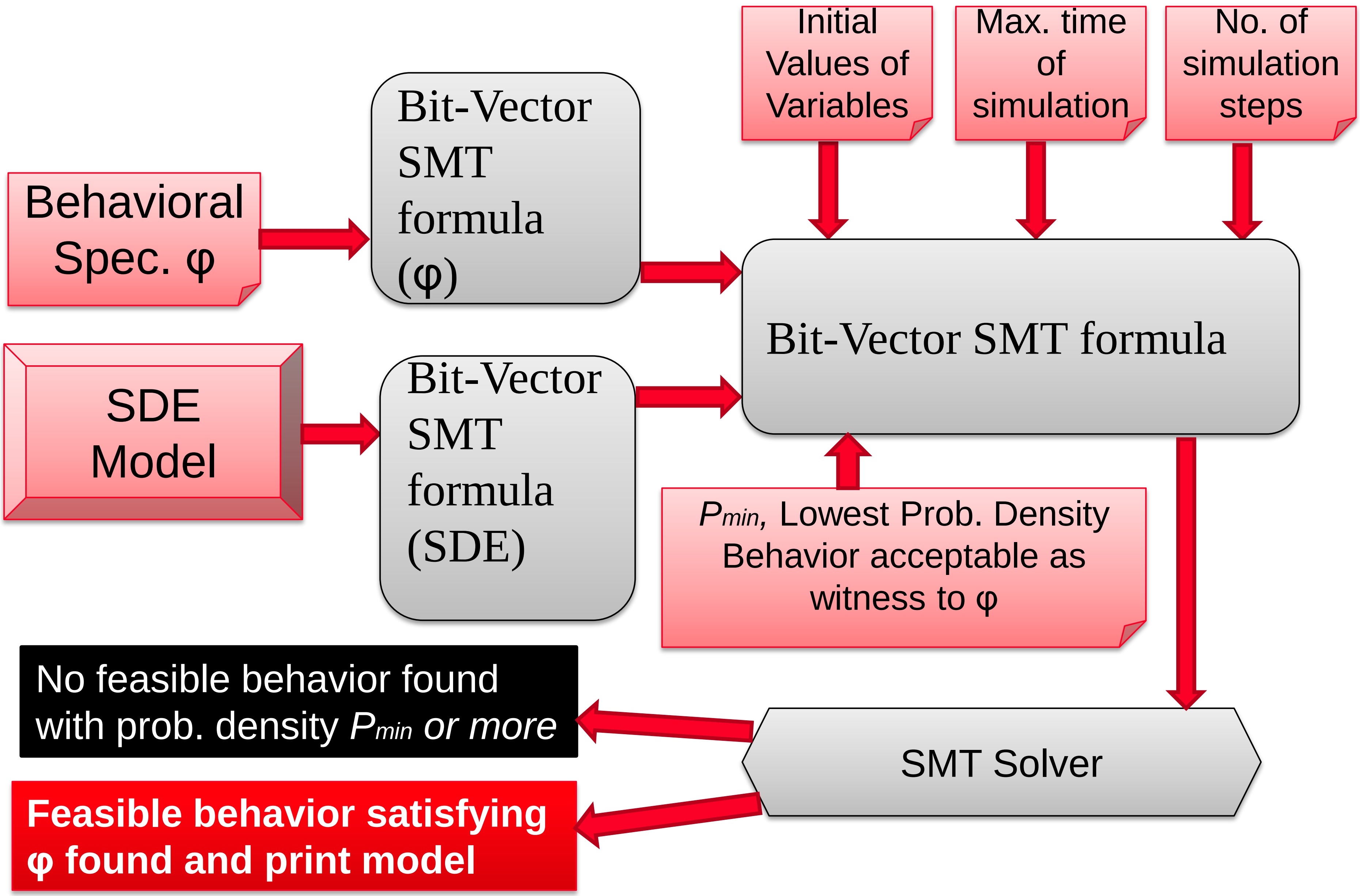
Our Approach – II (RESERCHE)



Our Approach – II (RESERCHE)



Our Approach – II (RESERCHE)



Our Approach - III

- SDE to Bit-Vector SMT formula

  - Discrete Difference Equation

$$\bigwedge_{k=0}^{m-1} \left(X_{t_{k+1}} = X_{t_k} + b\,(t_k, X_{t_k})(t_{k+1} - t_k) + v\,(t_k, X_{t_k})\ \Delta W_{t_{k+1} \leftrightarrow t_k} \right)$$

  - ΔW is increment in Brownian motion
    – Normally distributed
    – with mean 0
    – and variance 1

  - $t_{k+1} - t_k$ is increment in time

Intuitive Argument

- Large values of Brownian motion increments

  ➔ smaller probability densities
- Small values of Brownian motion increments

  ➔ large probability densities

A behavior has high probability density if it corresponds to small values of Brownian motion increments

Intuitive Argument

- Large values of Brownian motion increments

  ➔ smaller probability densities
- Small values of Brownian motion increments

  ➔ large probability densities

A behavior has high probability density if it corresponds to small values of Brownian motion increments

Ask a decision procedure if there is a behavior satisfying the given specification with small values of Brownian motion increments.

Digging Deeper…

- **The probability density of observing the value X_{t_1} after t_1 time:**

$$P(X_{t_1}|X_{t_0}) = P\left(W_{t_1} - W_{t_0} = \widehat{W_{t_1}} - \widehat{W_{t_0}}|X_{t_0}\right)$$
$$= P\left(W_{t_1} - W_{t_0} = \widehat{W_{t_1}} - \widehat{W_{t_0}}|W_{t_0} = \widehat{W_{t_0}}\right)$$

$$= \frac{1}{\sqrt{2\pi}(t_1 - t_0)} \; e^{-\left(\frac{|\widehat{W_{t_1}} - \widehat{W_{t_0}}|)^2}{2(t_1 - t_0)}\right)}$$

- **Our results rely on**
  - **the independence of increments of Brownian Motions, and**
  - **their Gaussian distribution.**

Digging Deeper… …

- We compute the probability density of observing the sequence of the observed discretized solution given the initial value:

$$
\begin{aligned}
&P(X_{t_0}, X_{t_1}, \ldots X_{t_m} | X_{t_0}) \\
&= P(X_{t_1}, \ldots X_{t_m} | X_{t_0}) P(X_{t_0} | X_{t_0}) \\
&= P(X_{t_1}, X_{t_2}, \ldots X_{t_m} | X_{t_0}, X_{t_1}) \quad Since,\ P(X_{t_0} | X_{t_0}) = 1 \\
&= P(X_{t_2}, \ldots X_{t_m} | X_{t_0}, X_{t_1}) P(X_{t_1} | X_{t_0}) \\
&= P(X_{t_2}, \ldots X_{t_m} | X_{t_1}) P(X_{t_1} | X_{t_0}) \ \ Since,\ P(X_{t_2} | X_{t_0}, X_{t_1}) = P(X_{t_2} | X_{t_1}) \\
&= P(X_{t_3}, \ldots X_{t_m} | X_{t_1}, X_{t_2}) P(X_{t_2} | X_{t_1}) P(X_{t_1} | X_{t_0}) \\
&= \ldots \qquad\qquad\qquad \ldots \qquad\qquad\qquad \ldots \\
&= P(X_{t_m} | X_{t_{m-1}}) \ \ \ldots \ \ P(X_{t_1} | X_{t_0}) \\
&= (2\pi)^{-m/2} \Delta^m e^{-\frac{m}{2\Delta}\left(\sum_{i=1}^{m}\left(\widehat{W_{t_i}} - \widehat{W_{t_{i-1}}}\right)^2\right)}
\end{aligned}
$$

Digging Deeper… …

- We compute the probability density of observing the sequence of the observed discretized solution given the initial value:

$$
\begin{aligned}
&P(X_{t_0}, X_{t_1}, \ldots X_{t_m} | X_{t_0}) \\
&= P(X_{t_1}, \ldots X_{t_m} | X_{t_0}) P(X_{t_0} | X_{t_0}) \\
&= P(X_{t_1}, X_{t_2}, \ldots X_{t_m} | X_{t_0}, X_{t_1}) \quad Since,\ P(X_{t_0} | X_{t_0}) = 1 \\
&= P(X_{t_2}, \ldots X_{t_m} | X_{t_0}, X_{t_1}) P(X_{t_1} | X_{t_0}) \\
&= P(X_{t_2}, \ldots X_{t_m} | X_{t_1}) P(X_{t_1} | X_{t_0}) \ \ Since,\ P(X_{t_2} | X_{t_0}, X_{t_1}) = P(X_{t_2} | X_{t_1}) \\
&= P(X_{t_3}, \ldots X_{t_m} | X_{t_1}, X_{t_2}) P(X_{t_2} | X_{t_1}) P(X_{t_1} | X_{t_0}) \\
&= \ldots \qquad\qquad \ldots \qquad\qquad \ldots \\
&= P(X_{t_m} | X_{t_{m-1}}) \ \ \ldots \ \ P(X_{t_1} | X_{t_0}) \\
&= (2\pi)^{-m/2} \Delta^m e^{-\frac{m}{2\Delta}\left(\sum_{i=1}^{m}\left(\widehat{W_{t_i}} - \widehat{W_{t_{i-1}}}\right)^2\right)}
\end{aligned}
$$

- We need to minimize the sum of squares of Brownian motion increments!

Concerns

- A key concern in discretizing SDE is the error introduced by sampling a continuous system and replacing a SDE with a discretized difference equation.

- The existence and uniqueness of SDE ensures that the solution of an infinitely discretized SDE is the solution of the continuous SDE.

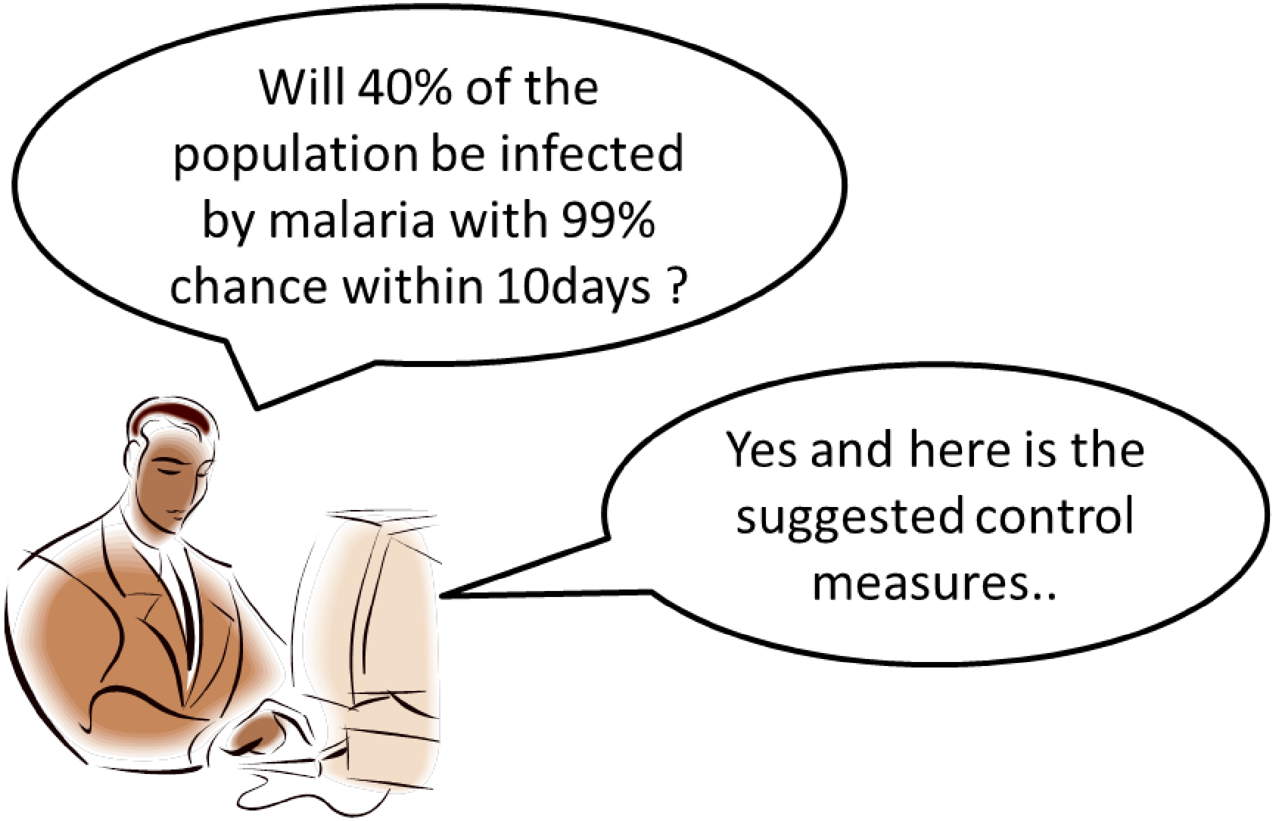
- What is sufficiently discretized?
  - 100, 1000, 10^4,…,10^100
  - Stochastic mean value theorem and rate of convergence for various drift and diffusion.

Discretizing SDE vs. ODE

- Does a point lie on the trajectory of an ODE?
  - Not answerable in general
  - Works in practice; Lipschitz Continuity
- Does a point lie on some path of a SDE?
  - Yes

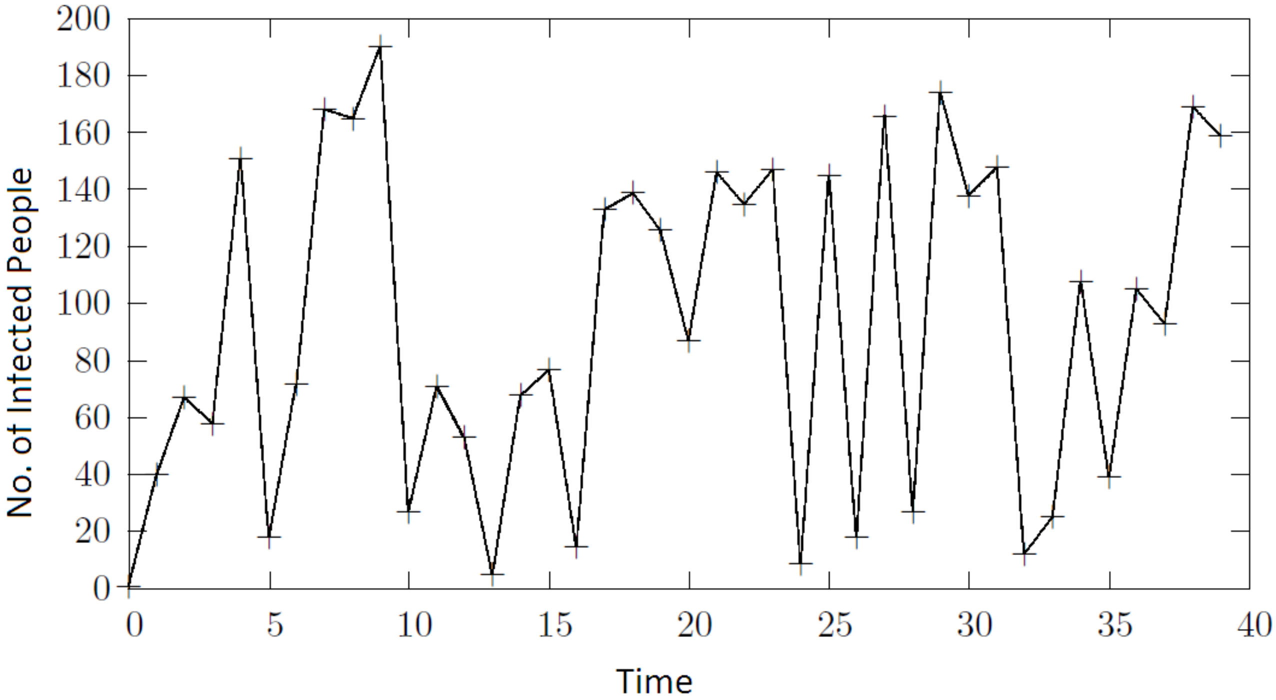
- Discretizing ODEs is fine with knowledge of Lipschitz constants
- Discretizing SDEs is fine with knowledge of Lipschitz constants for drift and diffusion

Case Study – High Level View



A high level view of our system: A person interacting with our system through web interface and our system is providing the instant results to the queries

Experiments - I

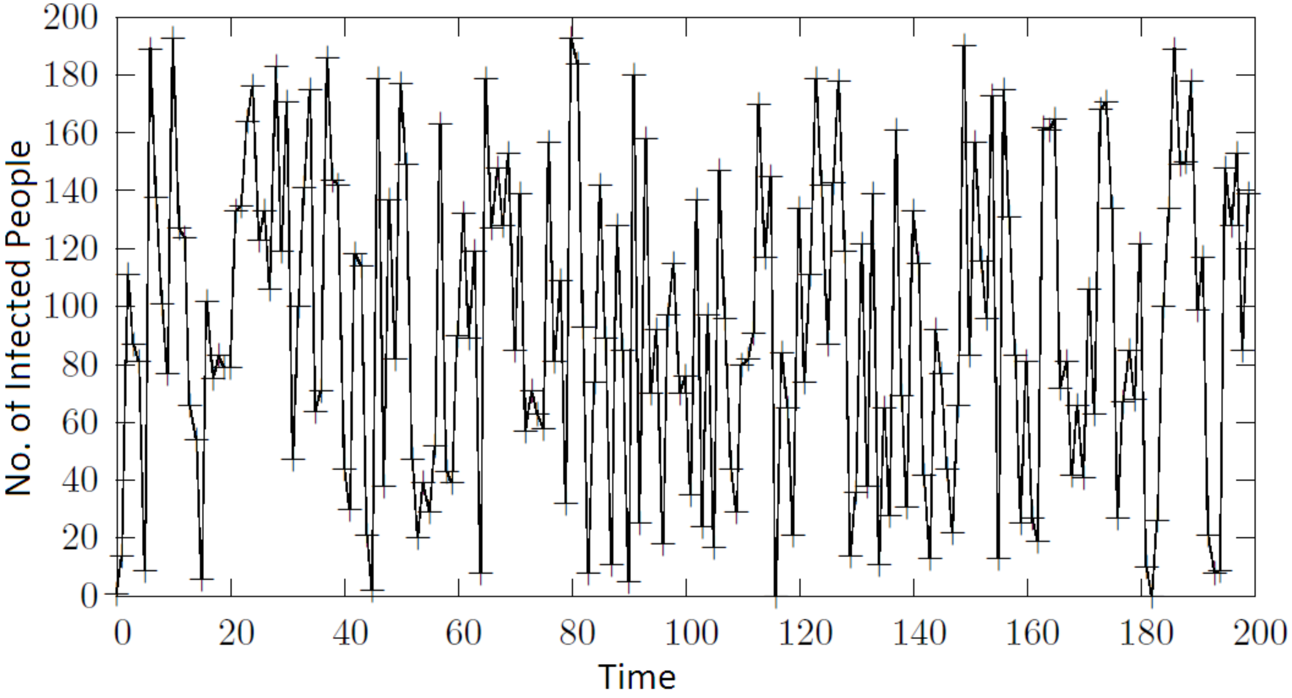


ln(prob. density) = -2.4e+008

SARS: The number of infected people vs. Time plot.

Above prob. density is calculated when the number of infected people is more than 150 and less than 200 at the end of 40 time steps. Initially, no of infected people = 1

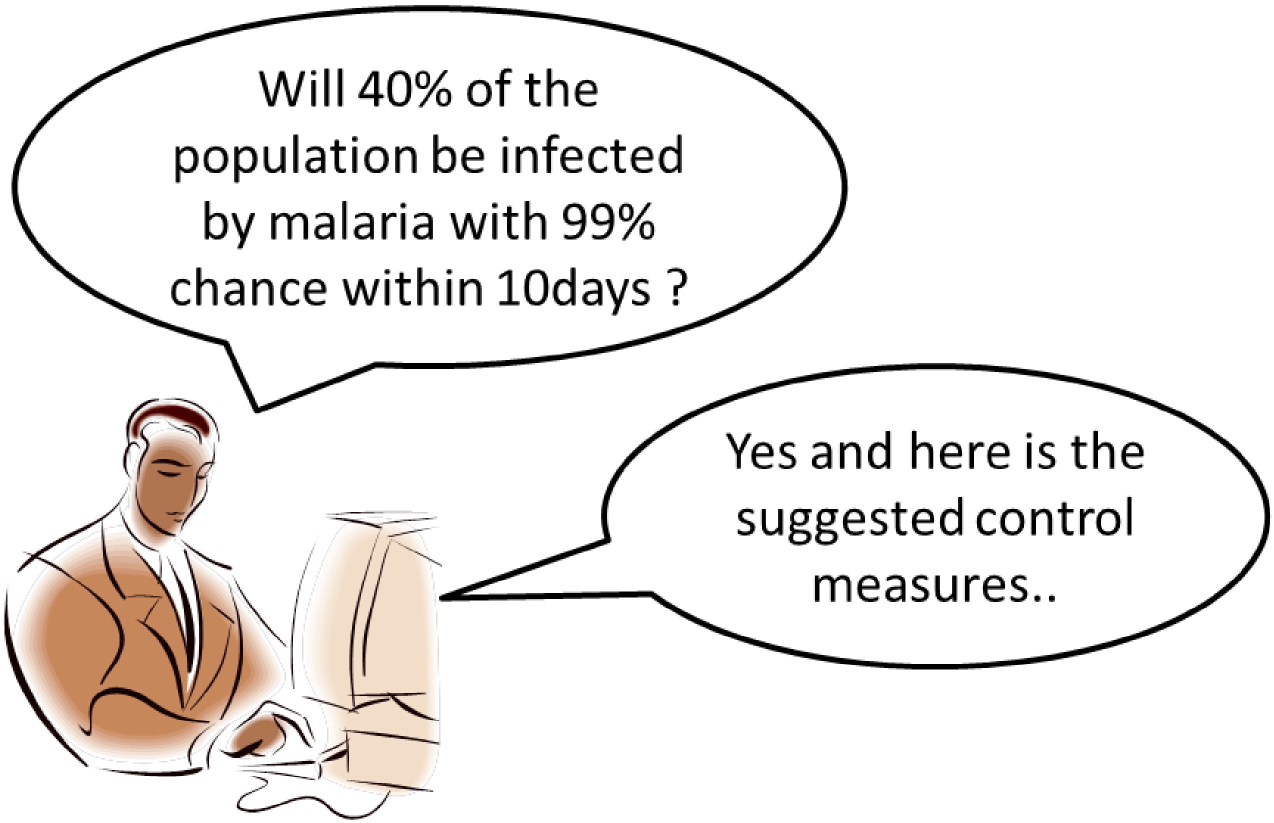
Experiments - II



SARS: The number of infected people vs. Time plot.

No. of time steps = 200

Case Study – High Level View



A high level view of our system: A person interacting with our system through web interface and our system is providing the instant results to the queries

Conclusion

- Algorithm for efficiently investigating rare behaviors in SDE models.
  - It avoids the computational costs associated with sampling
    - by searching for trajectories from the model that satisfy a given behavioral specification.
  - Only generates trajectories that exhibit the behavior.
- Our method takes advantage of the efficiency and power of the modern SMT-solvers.

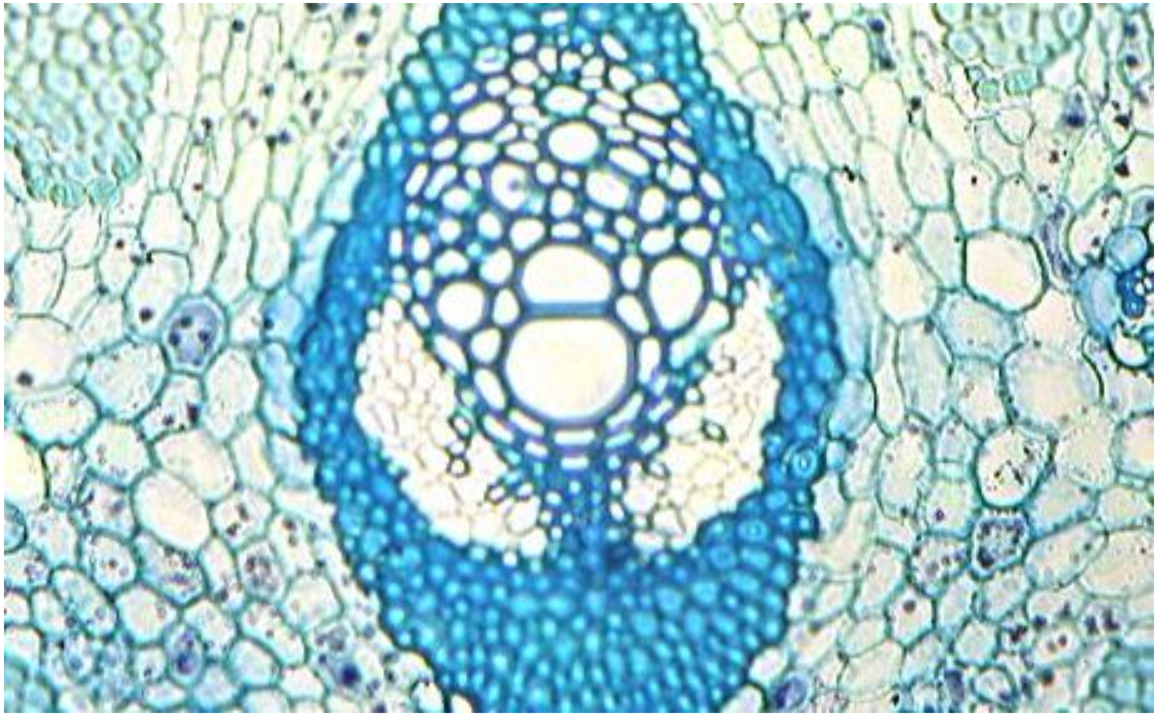
Future Work

- Studying the use of decision procedures to analyze closed form solutions to SDEs.
  - Many practical applications require the study of the system where one component is SDE and the other component is a finite state controller.
    - Biologically important cyber-physical systems like artificial pancreas.
- Specialized decision procedures
  - Sum of Squares
  - Grobner Bases
  - Nonlinear SMT
- Extensive parallelization.

Future Work

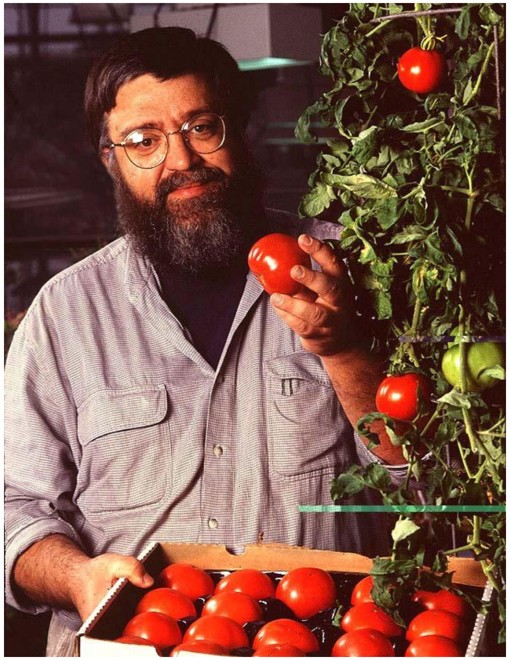
- Integrated Circuit Performance
- Artificial Pancreata
- Computational Finance
- Autonomous Vehicles

Synthetic Biology



Genetic Similarity – Phenotypic Diversity

Synthetic Biology



FlavrSavr – Evolving away from synthetic target